An Evaluation of Convolutional Neural Networks for Lithological Mapping Based on Hyperspectral Images

Ziye Wang and Renguang Zuo

Abstract—**Hyperspectral remote sensing images are characterized by nanoscale spectral resolution and hundreds of continuous spectral bands, dominating significantly in geological applications ranging from lithological mapping to mineral exploration. A major challenge lies in how to incorporate spectral and spatial information, therefore promoting classification performance for detecting closely resembling and mixed minerals and lithologies. Recent advances in deep learning techniques have facilitated the application of hyperspectral images in geological studies, especially experts at handling high-dimensional data with strong neighboring correlation. As a result, this study focuses on the evaluation of deep learning algorithms for lithological mapping based on hyperspectral images and further provides guidance on mineral exploration. Four deep convolutional neural networks (CNNs), including 1-D CNN, 2-D CNN, 3-D CNN, and a hybrid of 1-D and 2-D CNN, were constructed for spectral, spatial, and spatial–spectral feature extraction. The proposed frameworks were verified through the case studies of lithological mapping to aid in prospecting rare metal deposits using Gaofen-5 hyperspectral images in the Cuonadong dome, Tibet, China. Lithological classification maps indicated that the dual-branch 1D–2D CNN yields better performance than others in both visual and quantitative comparisons, owing to the support of joint spatial–spectral feature learning. An overall classification accuracy of up to 97.4% further illustrates the feasibility of CNN models for lithological mapping based on hyperspectral images, which provides a realizable and promising approach for mineral exploration by mapping specific lithologies.**

Index Terms—**Convolutional neural networks (CNNs), Gaofen-5 (GF-5), hyperspectral image, lithological mapping.**

I. INTRODUCTION

HYPERSPECTRAL remote sensing refers to a technology that images targets in tens to hundreds of spectral bands using a hyperspectral sensor [1]. The combination of imagery and spectra increases the capability of detecting more detailed surface materials than multispectral images, thus has been extensively used in the fields of agriculture [2], military [3], environmental monitoring [4], geological mapping, and mineral exploration [5], [6], [7]. Taking geological mapping as an example, minerals have specific spectral absorption and texture properties in the visible- and near-infrared (VNIR) and shortwave infrared (SWIR) spectral ranges, owing to various compositions in the crystal structure, physical, and chemical patterns [8]. Therefore, targeted lithologies and surroundings can be distinguished and recognized using specific spectral diagnostic features [9], [10].

Hyperspectral remote sensing images can be regarded as a 3-D data cube, including spectral and spatial information [11]. The spectral domain records the spectral reflectance of targets with a change in wavelength; the spatial domain is a 2-D image that reflects the position of each pixel and its neighborhoods, providing complementary information for the extraction of spectral features [12]. Hyperspectral image classification is fundamentally implemented by assigning a unique label to each pixel or image [13], [14], which can be divided into two categories: feature extraction approaches [15], such as principal component analysis (PCA) [16], independent component analysis [17], spectral feature fitting [18], minimum noise fraction transformation [19], manifold learning [20], linear discriminant analysis [21], band selection [22], and metric learning [23]; classification approaches, such as k-means [24], spectral angle mapper [25], support vector machine [26], random forest [27], subspace learning [28], and more recent deep learning algorithms, such as convolutional neural network (CNN) [29], graph neural network [30], recurrent neural network [31], deep belief networks [32], stacked autoencoders (SAE) [33], and generative adversarial networks [34].

The pixelwise approaches regard hyperspectral images as an ensemble of spectral information that ignores the spatial structure of adjacent pixels, potentially resulting in isolated and discrete classifications [35]. Alternatively, other methods are based on spatial features, such as the Markov random field [36], extended morphological profiles [37], gray-level co-occurrence matrix [38], and Gabor filters [39]. Spatial information from adjacent pixels is extracted by such kind of methods to improve classification accuracy. However, these patchwise approaches break the original spectral sequence without any concerns about the spectral features of hyperspectral images, which are equally important for classification.

Each pixel and its neighborhoods in a hyperspectral image are locally spatially correlated and composed of continuous spectral curves in a sequence. Therefore, the joint integration of spatial–spectral features within neighboring regions as the input can lead to learning more discriminating features than spatialwise or spectralwise classification [40], [41], [42], [43].

Manuscript received 4 December 2023; revised 24 January 2024; accepted 25 February 2024. Date of publication 1 March 2024; date of current version 18 March 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 42102332 and Grant 42372344, in part by China Postdoctoral Science Foundation under Grant 2021M692988, and in part by the Natural Science Foundation of Hubei Province (China) under Grant 2023AFA001. *(Corresponding author: Renguang Zuo.)*

The authors are with the State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China (e-mail: ziyewang@cug.edu.cn; zrguang@cug.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3372138

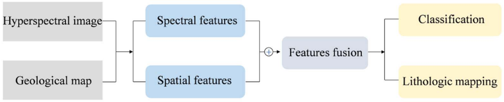


Fig. 1. Brief framework of lithological mapping based on hyperspectral images.

As an end-to-end deep learning algorithm that can perform representation learning and classification tasks simultaneously, CNNs are designed to learn complex nonlinear distribution features in images by automatically extracting high-level spatial features, which are especially suitable for dealing with high-dimensional hyperspectral images with strong neighboring correlations [44], [45]. In this aspect, 3-D CNN [46], fully convolutional network [47], convolutional capsule network [48], and multiple deep learning or dual-channel frameworks, such as ConvLSTM [49], [50], 1D–2D CNN [51], 3D–2D CNN [52], [53], convolutional autoencoder [54], graph convolutional network [55], CNN-local discriminant embedding [56], CNN-SAE [57], CNN-transformer learning [58], and joint attention network [59], have been introduced for spectral–spatial hyperspectral classification and achieved state-of-the-art performance.

The combined use of CNNs and hyperspectral data provides the potential to assist in geological mapping involving mineral exploration [60], [61], [62], [63]. On this account, this study focused on using CNNs to exploit hyperspectral data for lithological mapping. Four CNN models, namely, 1-D CNN, 2-D CNN, 3-D CNN, and 1D–2D CNN, were constructed for spectral, spatial, and joint spatial–spectral feature extraction, respectively. The 1D–2D CNN is a dual-branch structure model (see Fig. 1) that incorporates a 1-D CNN with a 2-D CNN, in which 1-D convolutions are employed to extract spectral features by exploring the contextual information between spectral bands, and 2-D convolutions are used to extract local spatial features and neighboring relationships. The four CNNs were demonstrated by a case study of lithological mapping in the Cuonadong dome, the northeastern Himalayan orogenic belt, based on Gaofen-5 (GF-5) hyperspectral images. Notably, mineralogical research supports the high potential of rare metal mineralization, such as Be, W, Li, and Sn in this area [64], [65], [66], [67], further promoting the significance of lithological mapping to aid in mineral exploration. The main contributions of this study include the following:

1) evaluating the feasibility of CNNs for lithological mapping based on GF-5 hyperspectral remote sensing images;
2) providing a realizable and valuable approach for mineral exploration by mapping specific lithological units.

II. METHODS

A. CNN Architecture Setting

CNN is a deep learning algorithm with multiple hidden layers based on the idea of local connections and weight sharing [68], [69]. A typical CNN architecture consists of input, convolutional, pooling, and fully connected layers [70]. The first three layers perform feature extraction, and the fourth layer maps the extracted features to the output and performs classification [71]. Especially, the convolutional layer extracts spatial dependencies in the input image through convolutional filters with kernels. Features ranging from low to high level are gradually and adaptively learned per layer. The shared weights and locally connected layers equipped in the convolutional operation significantly increase the model efficiency by reducing the number of parameters. The pooling layer is responsible for the downsampling operation and reduces the spatial size of the convolved features. This process decreases computational complexity, therefore reducing overfitting through dimensionality reduction [72]. The last fully connected layer is also called the output layer, designed to make predictions using the SoftMax activation function, such as the probabilities of each class in lithological mapping tasks. Depending on the dimensions of the convolution kernel, CNNs can be divided into the following three types.

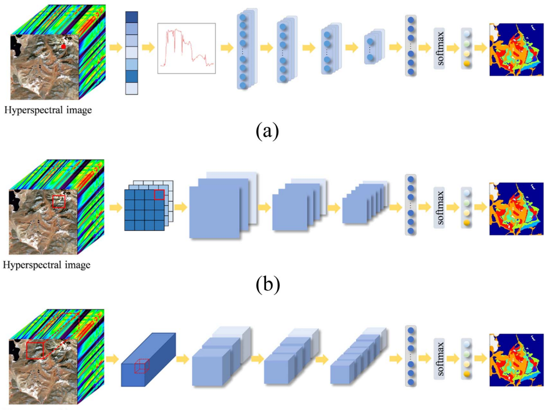


Fig. 2. Network architectures of (a) 1-D CNN, (b) 2-D CNN, and (c) 3-D CNN.

1) One-Dimensional CNN: Each pixel in a hyperspectral image is composed of a continuous spectral vector, which can be regarded as a sequence-based data structure. The 1-D CNN is specifically designed to handle 1-D data by performing pixelwise classification using 1-D convolutional kernels/filters [73]. Benefiting from a few computational complexities and parameters, 1-D CNN is easier to implement by capturing the contextual spectral information of each pixel within a sequence. The 1-D CNN used in this study shares the basic structure of a common CNN, including three convolution layers with 3 × 1 kernels, three max-pooling layers with 2 × 1 kernels, two batch normalization layers, a SoftMax function as a fully connected layer, and a ReLU as the activation function [see Fig. 2(a)]. Here, the normalization operation normalizes the input data in each batch to prevent gradient explosions as well as accelerates the training process.

2) Two-Dimensional CNN: 1-D CNN usually reshapes the hyperspectral image into a vector and loses the spatial features. As an alternative, 2-D convolution is introduced to extract spatial neighboring information by taking the data patch delimited in a given window around each pixel. In a 2-D CNN, the kernel slides

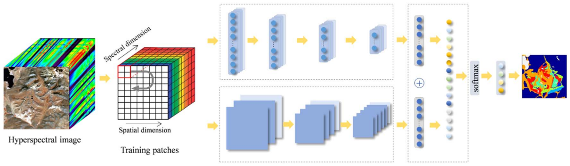


Fig. 3. Network architecture of 1D–2D CNN.

along two dimensions on the image, as shown in Fig. 2(b), which considers the local spatial correlations among the neighboring pixels [74]. The 2-D CNN used in this study comprised three convolutional layers, followed by two batch normalizations, three max-pooling, and two fully connected layers (dense layers). Each convolutional layer with a size of 3 × 3 was utilized to extract spatial information. Then, max-pooling layers with 2 × 2 kernels helped to reduce the data dimensions and the number of parameters. The ReLU activation function was also selected to improve the nonlinear mapping ability of the network [see Fig. 2(b)].

3) Three-Dimensional CNN: The 3-D CNN is a variant of CNNs that can extract features directly in both spectral and spatial dimensions through 3-D convolution operations [46], [75]. Fig. 2(c) illustrates the architecture of the 3-D CNN used in this study. In a 3-D convolution operation, the 3-D kernel slides in three directions, and the input cube is convolved with three 3-D kernels (3 × 3 × 3), inserting by three 3-D pooling operations with a kernel size of 2 × 2 × 2 and two fully connected layers. The 3-D CNN preserves the original 3-D data structure of hyperspectral images, thus contributing to a higher classification performance owing to the simultaneous learning of spectral–spatial features.

B. 1D–2D CNN Architecture Setting

The 1D–2D CNN architecture involves using 2-D convolutional layers for spatial feature representation combined with 1-D convolutional layers to support spectral sequence prediction, avoiding the loss of either spectral or spatial information [51], [76]. The 1D–2D CNN is a dual-branch structure comprising three parts: spectral information extraction by a 1-D CNN, spatial information extraction by a 2-D CNN, and integration of spatial–spectral information by a fusion operation. Fig. 3 describes how a hybrid 1D–2D CNN work for lithological mapping. The networks of 1-D and 2-D branches remain the same as those described above. The 1D–2D CNN starts with an input layer that is regarded as the combination of 1-D vectors corresponding to each pixel and spatial neighborhoods centered around the pixel. Then, spectral and spatial features were extracted using 1-D and 2-D convolutional kernels, respectively. Subsequently, the extracted spectral and spatial features were flattened into a 1-D vector and connected end-to-end. The joint features were finally fed into the fully connected layer where the SoftMax function serves as the classifier.

Hyperspectral images are presented in the form of 3-D data cubes. In general, the input of a 1-D CNN is a 1-D vector corresponding to each pixel in a hyperspectral image. The input of a 2-D CNN is a 2-D spatial patch centered on a certain pixel. And a 3-D CNN takes 3-D cubes within the neighborhood of each pixel as inputs to the network. The detailed architectures of these three networks are presented in Tables I–III, respectively.

TABLE I
NETWORK ARCHITECTURES OF 1-D CNN

Layer name	Number	Filter size	Feature dimensions
Input	/	/	1 × 30
Convolution layer	16	3 × 1	30 × 16
Batch normalization	/	/	30 × 16
Max-pooling layer	/	2 × 1	15 × 16
Convolution layer	32	3 × 1	15 × 32
Batch normalization	/	/	15 × 32
Max-pooling layer	/	2 × 1	7 × 32
Convolution layer	64	3 × 1	7 × 64
Max-pooling layer	/	2 × 1	3 × 64
Flatten	/	/	192
Dense	/	/	64
Dense	/	/	8
SoftMax	/	/	8
Output	/	/	8
Total parameters	20904		

TABLE II
NETWORK ARCHITECTURES OF 2-D CNN

Layer name	Number	Filter size	Feature dimensions
Input	/	/	21 × 21 × 30
Convolution layer	16	3 × 3	21 × 21 × 16
Batch normalization	/	/	11 × 11 × 16
Max-pooling layer	/	2 × 2	11 × 11 × 16
Convolution layer	32	3 × 3	11 × 11 × 32
Batch normalization			11 × 11 × 32
Max-pooling layer	/	2 × 2	6 × 6 × 32
Convolution layer	64	3 × 3	6 × 6 × 64
Max-pooling layer	/	2 × 2	3 × 3 × 64
Flatten	/	/	2304
Dense	/	/	64
Dense	/	/	8
SoftMax	/	/	8
Output	/	/	8
Total parameters	65112		

TABLE III
NETWORK ARCHITECTURES OF 3-D CNN

Layer name	Number	Filter size	Feature dimensions
Input	/	/	21 × 21 × 30
Convolution layer	16	3 × 3 × 3	21 × 21 × 30 × 16 □
Batch normalization	/	/	21 × 21 × 30 × 16
Max-pooling layer	/	2 × 2 × 2	10 × 10 × 15 × 16
Convolution layer	32	3 × 3 × 3	10 × 10 × 15 × 32
Batch normalization	/	/	10 × 10 × 15 × 32
Max-pooling layer	/	2 × 2 × 2	5 × 5 × 7 × 32
Convolution layer	64	3 × 3 × 3	5 × 5 × 7 × 64
Max-pooling layer	/	2 × 2 × 2	2 × 2 × 3 × 64
Flatten	/	/	768
Dense	/	/	64
Dense	/	/	8
SoftMax	/	/	8
Output	/	/	8
Total parameters	119,592		

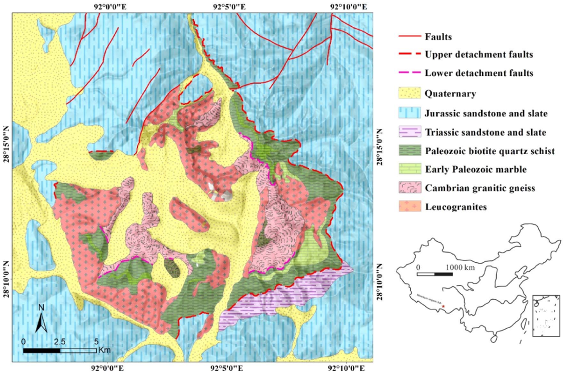


Fig. 4. Simplified geological map of the Cuonadong dome, northeast Tibet, China [79].

C. Metrics

Three quantitative metrics, such as classification accuracy, overall accuracy, and kappa coefficient, were employed to measure the performance of the four networks. The classification accuracy is defined as the ratio of each category being correctly classified. The overall accuracy is the percentage of total sample correctly classified. The kappa coefficient is a robust index used for consistency testing, that is, whether the predicted results are consistent with the actual classification results [77]. These metrics are essentially calculated based on a confusion matrix. A confusion matrix is a visual representation of the performance of a model, which summarizes the percentages of categories being correctly classified and misclassified into others.

III. STUDY AREA AND DATASET

A. Geological Background

The study area, Cuonadong dome, is located in the eastern part of the Himalayan orogenic belt and belongs to the famous Zhaxikang ore concentration region [78]. Fig. 4 describes the simplified geological map of the Cuonadong dome [79]. This area is a gneiss dome cut by E–W and N–W faults, which is basically covered with sandstone, slate, granitic gneiss, marble, biotite quartz schist, and leucogranite [80], [81]. Abundant rare metal mineralization has been discovered in this area, such as Be, Nb, Ta, Sn, Pb, W, U, and Li. For example, Li et al. [64] reported the W–Sn and Be–Rb rare metal mineralization hosted in skarn marble, pointing out that large-scale Be–W–Sn deposits are closely related to Himalayan leucogranites in the Cuonadong dome. Xia et al. [82] recommended that the exploration of rare metal deposits should be carried out around granite pegmatite belts in which the possible targets are skarn and leucogranites. Cao et al. [67] systematically analyzed stream sediment geochemical data in the Himalayan orogenic belt, indicating positive rare metal anomalies associated with leucogranites compared with the upper continental crust. However, an average altitude of 5000 m and poor environment limits the conduct of field geological survey. Consequently, delineating the distribution of lithological classes using remote sensing technology is of great significance to assist in discovering rare metal deposits in the study area.

The main mineral compositions of leucogranite include quartz (35%–42%), orthoclase (15%–20%), plagioclase (30%–40%), muscovite (7%–10%), and biotite (1%–2%), with particle diameters ranging from 1 mm to 1 cm [83]. These minerals are in response to various spectral absorption that can be discriminated via the diagnostic signatures in the VNIR and SWIR regions (0.4–2.5 μm). For example, research has confirmed that granodiorite shows an absorption peak from 0.45–0.52 μm due to ferromagnetic minerals. In contrast, absorption peak positions of carbonate minerals (such as calcite and dolomite) occur from 2.30–2.39 μm [5], [6], [84]. Spatially contiguous spectra captured by hyperspectral images are more sensitive to subtle variations in spectral absorption features, providing the potential to quantitatively identify mineral categories and delineate the spatial distribution of lithological units [62].

B. Hyperspectral Images

The GF-5 is the latest comprehensive hyperspectral observation satellite launched by China. The 30-m-spatial-resolution advanced hyperspectral imager installed on GF-5 provides 330 spectral bands ranging from 0.4 to 2.5 μm. What is more, GF-5 poses more advantages over a multispectral image in the ability to distinguish similar minerals owing to 150 VNIR bands with a spectral resolution of 5 nm and 180 SWIR bands with a spectral resolution of 10 nm [85].

The GF-5 hyperspectral image used in this study is a level 1 product acquired on Oct. 13, 2019, covering an area of approximately 600 km^2. Preprocessing of the GF-5 data comprised band selection, radiometric calibration, atmospheric correction, noise reduction, and geometric correction. Band selection was attempted to remove 35 repeated and affected bands due to water vapor absorption. Radiometric calibration and atmospheric correction were then implemented to convert raw digital images from satellites into actual physical surface reflectance and eliminate scattering and absorption effects from the atmosphere. These two operations were primarily carried out using the FLAASH module in the ENVI software [86]. Regarding the geometric correction, it corrects the geometric distortions and positional errors with the help of a reference coordinate system, for example, the Landsat-8 imagery used in this study [87]. Finally, 815 × 786 image with 30 m resolution and 295 hyperspectral bands was adopted for lithological mapping in the study area (see Fig. 5).

Seven lithological units, including Jurassic sandstone and slate (class 1), early Paleozoic marble (class 2), Triassic sandstone and slate (class 3), Cambrian granitic gneiss (class 4), Paleozoic biotite quartz schist (class 5), Quaternary strata (class 6), Himalayan leucogranites (class 7), and an additional class representing lakes and snow (class 8), were designed to identify in the study area. Table IV lists the areas and mineral compositions of each category [79]. The spectral curves of these lithological units are shown in Fig. 6. It was observed that the seven lithological units remain consistent in the overall spectral shape while varying slightly in absorption and reflection peaks, implying strong band correlation. More detailly, these seven lithological units present similar reflection features near 0.60 and 0.792 μm in VNIR and 1.249, 1.418, and 1.839 μm in SWIR, and

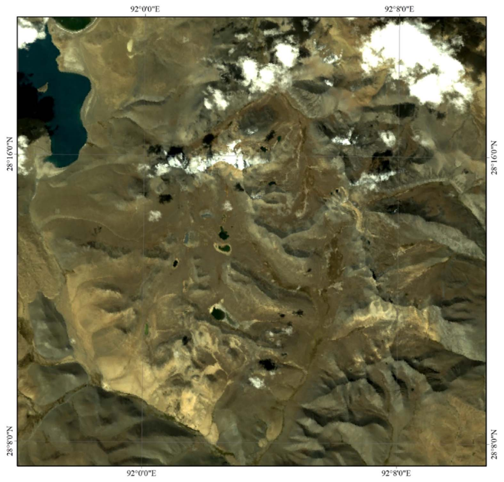
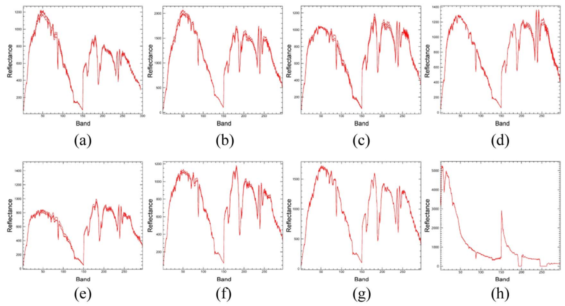


Fig. 5. GF-5 hyperspectral image of the study area.

TABLE IV
AREAS AND MINERAL COMPOSITIONS OF EACH LITHOLOGICAL UNIT IN THE STUDY AREA [79]

Class	Lithological unit	Mineral composition	Areas (km²)
1	Jurassic sandstone and slate	Mudstone, metamorphic sandstone, or carbonaceous calcareous slate	286.6
2	Early Paleozoic marble	Calcite, dolomite, muscovite, and quartz	11.5
3	Triassic sandstone and slate	Feldspar, shale, slate, quartz sandstone, and sandy slate	13.0
4	Cambrian granitic gneiss	Potassium feldspar, quartz, and biotite	29.5
5	Paleozoic biotite quartz schist	Quartz, biotite, plagioclase, garnet	47.8
6	Quaternary	Gravel, sand, and clay	125.9
7	Himalaya Leucogranite	Quartz, potash feldspar, plagioclase, muscovite, biotite, tourmaline	62.8
8	Lakes and snow	-	26.2



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 6. Spectral curves of lithological units in the study area. (a) Jurassic sandstone and slate. (b) Early Paleozoic marble. (c) Triassic sandstone and slate. (d) Cambrian granitic gneiss. (e) Paleozoic biotite quartz schist. (f) Quaternary. (g) Himalaya leucogranite. (h) Lakes and snow.

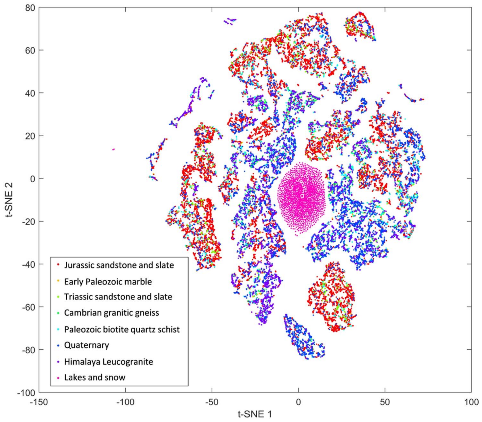


Fig. 7. Separability visualization of the hyperspectral image obtained by t-SNE.

absorption features near 0.771, 1.334, and 1.713 μm. The reason can be attributed to mixture spectra caused by intense weathering and erosion. Triassic sandstone and slate exhibit discernable spectral characteristics during 1.839–2.0 μm; Paleozoic biotite quartz schist is distinguishable from others around 1.206 μm. As the same silicate minerals, Himalaya leucogranite can be identified by the absorption feature from 2.0 to 2.4 μm. The premise of hyperspectral image classification is based on the separability of spectral features into different classes. Fig. 7 visualizes the separability of the GF-5 hyperspectral image using the t-distributed stochastic neighbor embedding (t-SNE) algorithm. t-SNE is a popular unsupervised cluster analysis method for visualization of high-dimensional data in a 2-D space [88]. Visual inspection indicates that the lithological units in the study area are separated by clustering, revealing that they are distinguishable from each other and can be identified using hyperspectral classification methods.

IV. EXPERIMENTAL SETTINGS

A. *Training Sample Preparation*

The size of the input hyperspectral data was $815 \times 786 \times D$, where D is the number of bands. A 1-D CNN typically reshapes the input into 1-D vectors and performs convolution operations in the spectral direction. A total of 10% of pixels in each class were randomly selected to train the 1-D CNN model. The training samples were labeled based on the reference geological map of the study area. The 2-D CNN creates patch-based samples by drawing a spatial neighborhood around each pixel as the center, thereby capturing the local spatial information contained in this patch. This process was implemented using the sliding window method at a stride of one, which also provides sufficient samples for model training. Theoretically speaking, a larger patch will bring better performance; however, an increased computational burden needs to be borne. The size of the neighboring region was empirically set to 21×21; in this context, more than 620 000 patches were produced. The 3-D CNN takes a cube as the input formed by every pixel along with its $21 \times 21 \times D$ neighbors, covering a 21×21 spatial extent and D spectral bands, making

TABLE V
PARAMETERS SETTING OF FOUR MODELS

Parameters	**1-D CNN**	**2-D CNN**	**3-D CNN**	**1D–2D CNN**
Epoch	300	300	300	300
Batch size	128	128	128	128
Learning rate	0.0001	0.0001	0.0001	0.0001
optimizer	Adam	Adam	Adam	Adam
Loss function	Cross Entropy	Cross Entropy	Cross Entropy	Cross Entropy
Window size	1	21	21	21

TABLE VI
ABLATION EXPERIMENT OF NETWORK ARCHITECTURE

Model	**Conv1**	**Pool1**	**Conv2**	**Pool2**	**Conv3**	**Pool3**	**Accuracy**
1	√	√					87.95%
2	√	√	√				93.55%
3	√	√	√	√			93.36%
4	√	√	√	√	√		93.97%
5	√	√	√	√	√	√	94.21%

Conv— convolutional layer; Pool— max-pooling layer

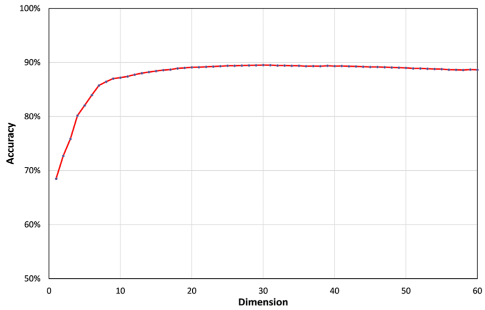


Fig. 8. Classification accuracy versus the hyperspectral dimension.

it possible to learn spectral–spatial information simultaneously. Regarding the 1D–2D CNN, the network and size of training samples were consistent with those of the 1-D and 2-D CNNs. The 10% of each class is randomly selected as a training sample set, of which 80% is for training and the remaining 20% is for verification.

B. Parameters Setting

Three parameters, including batch size, learning rate, and number of training epochs, need to be optimized for the CNN architectures [89]. The number of epochs refers to the number of times a learning process passes through the entire training dataset. The batch size defines the total number of training examples that work through a single batch, where the training dataset is usually divided into one or more batches for better computational efficiency. The learning rate is an important parameter determining whether a network can converge to the global minimum, which is set to control the apportioned error of a model [90]. These parameters are commonly determined via a grid search method optimized by the cross-entropy function. Table V provides the detailed parameter settings for the four networks.

C. Dimension Reduction

It was worth noting that hyperspectral bands should be dimensionally diminished to remove spectral redundancy. Fig. 8 presents how the spectral dimension influences the classification accuracy under different components, which was evaluated using the PCA. The model obtained the highest performance when the number of spectral bands was set to 30. As a result, the GF-5 hyperspectral image was reduced to 30 dimensions by PCA for the latter case studies.

D. Ablation Experiment

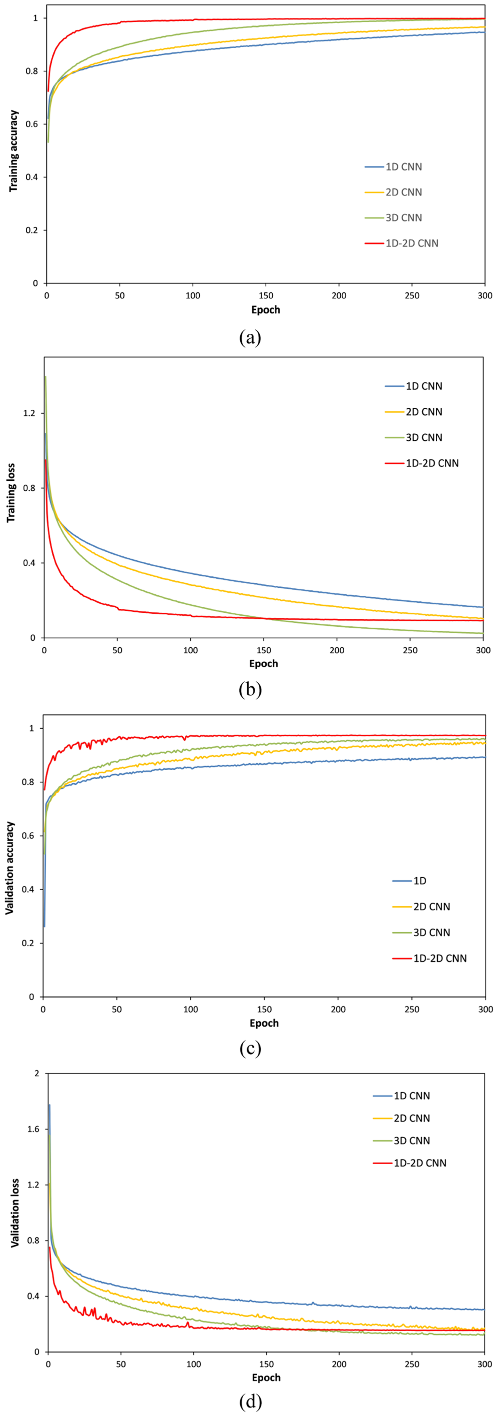
Understanding the influence of each network structure is crucial. An ablation experiment here was carried out to optimize the proposed network architectures. The ablation experiment is a set of studies to find the best-suited network by removing the parts of model components [91]. Table VI lists the classification accuracies of 2-D CNN with various architectures. The ablation study illustrates that the proposed network architecture reaches the best performance, on the other side, also provides a way to measure the contribution of each convolutional/pooling layer to the overall model.

E. Convergence Curves

The convergence curves of the epochs versus loss function/accuracy (see Fig. 9) remained steady after 300 epochs, suggesting that all models were well trained by the training sets. 1D–2D CNN converges faster in the training samples, whereas slower in the validation samples [see Fig. 9(a) and (c)], probably owing to a deeper network and more parameters needed to learn. In the case of the 3-D CNN, low loss values were obtained benefit from simpler network architecture [see Fig. 9(b) and (d)]. However, the use of hybrid CNNs reduces the computational complexity compared with 3-D CNN, thus outputting higher training accuracy.

V. RESULTS

The construction of CNNs' models and classifications of lithological units were carried out based on the TensorFlow platform. The GF-5 hyperspectral image was first segmented into patches of the same size using the sliding window method. All patches were then input into the trained models by coordinate order and output the corresponding categories. Fig. 10 shows the lithological classification maps obtained using 1-D, 2-D, 3-D, and 1D–2D CNN. The 1-D CNN [see Fig. 10(a)] suffers from salt pepper noise, resulting in poor performance visually. The 2-D CNN [see Fig. 10(b)] overcomes this drawback by learning the spatial correlation of each spectral dimension. The 3-D [see Fig. 10(c)] and 1D–2D [see Fig. 10(d)] CNNs achieved the best visual classification quality with fewer noisy and misclassified pixels than the 1-D and 2-D CNNs. The 3-D CNN performs roughly equivalently to the 1D–2D CNN



(a)

(b)

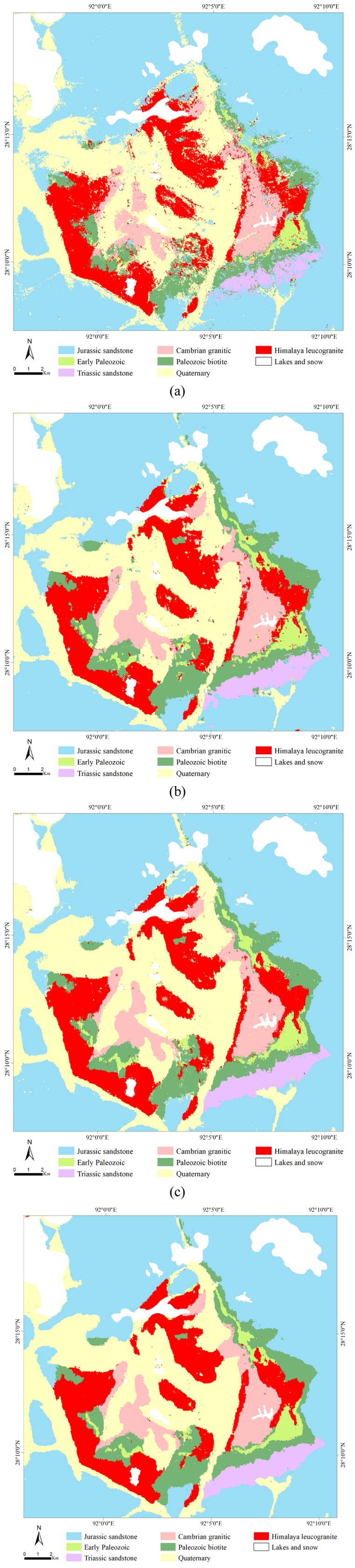
(c)

(d)

Fig. 9. Convergence curves of model training and validation. (a) Epoch versus training accuracy. (b) Epoch versus training loss. (c) Epoch versus validation accuracy. (d) Epoch versus validation loss.

due to the support of spatial–spectral feature learning, whereas the joint network suppresses more noise, leading to smoother boundaries and local details [see Fig. 11] compared with the direct extraction of spatial and spectral information by the 3-D CNN.

In addition to visual comparison, the confusion matrices of the four networks [see Fig. 12] also reveal that the 1D–2D CNN



(a)

(b)

(c)

(d)

Fig. 10. Lithological classification maps obtained by (a) 1-D CNN, (b) 2-D CNN, (c) 3-D CNN, and (d) 1D–2D CNN.

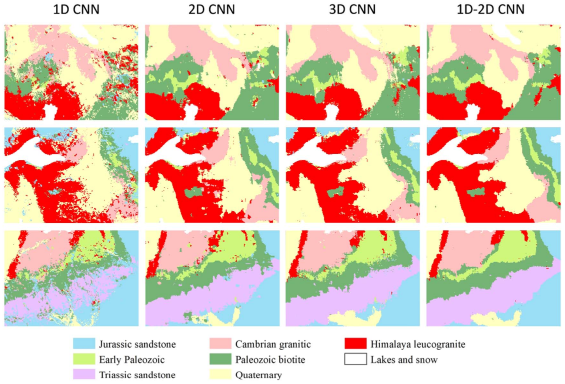


Fig. 11. Zoom of classification maps obtained by the four networks.

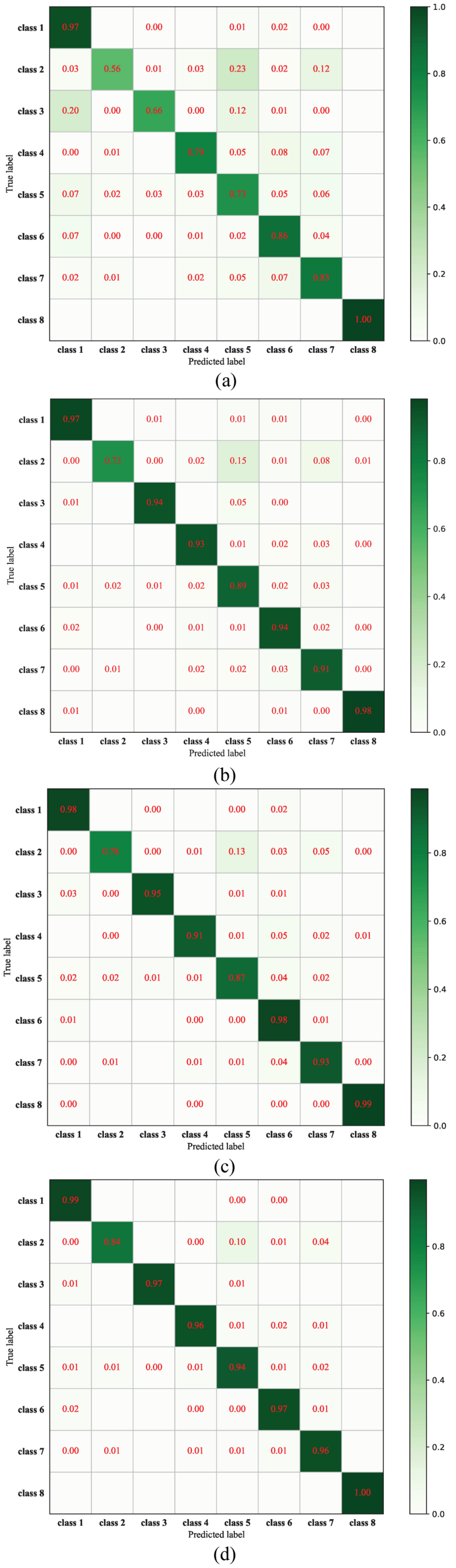
TABLE VII
QUANTITATIVE CLASSIFICATION COMPARISONS OF THE FOUR CASE STUDIES

Class	Lithological unit	1-D CNN	2-D CNN	3-D CNN	1D–2D CNN
1	Jurassic sandstone and slate	96.70%	96.98%	97.87%	99.10%
2	Early Paleozoic marble	55.82%	72.74%	77.57%	84.38%
3	Triassic sandstone and slate	66.00%	93.93%	94.64%	97.12%
4	Cambrian granitic gneiss	78.83%	93.35%	91.43%	95.86%
5	Paleozoic biotite quartz schist	73.07%	89.02%	87.25%	93.66%
6	Quaternary	86.17%	93.76%	97.68%	96.55%
7	Himalaya Leucogranite	82.85%	90.79%	92.92%	96.42%
	Overall accuracy	89.21%	94.21%	95.71%	97.39%
	Kappa	0.86	0.92	0.94	0.97
	Runtime (s)	660	839	6705	3490

outperforms others both in classification accuracy, overall accuracy, and kappa coefficient, agreeing with the aforementioned visual comparison. Two factors make 1D–2D CNN significantly improve the classification accuracy of each lithological unit. On the one hand, 1D–2D CNN makes full use of spectral and spatial features in hyperspectral data; on the other hand, the dual-branch structure integrates the two processes into a unified network, reducing the computational complexity that enables learning more discriminant information for lithological mapping.

VI. DISCUSSION

Table VII reports the quantitative comparison of the four networks, presenting varying performance for each lithological unit. In general, the joint extraction of spatial and spectral



	class 1	class 2	class 3	class 4	class 5	class 6	class 7	class 8
class 1	0.97		0.00		0.01	0.02	0.00	
class 2	0.03	0.56	0.01	0.03	0.23	0.02	0.12	
class 3	0.20	0.00	0.66	0.00	0.12	0.01	0.00	
class 4	0.00	0.01		0.79	0.05	0.08	0.07	
class 5	0.07	0.02	0.03	0.03	0.73	0.05	0.06	
class 6	0.07	0.00	0.00	0.01	0.02	0.86	0.04	
class 7	0.02	0.01		0.02	0.05	0.07	0.83	
class 8								1.00

(a)

	class 1	class 2	class 3	class 4	class 5	class 6	class 7	class 8
class 1	0.97		0.01		0.01	0.01		0.00
class 2	0.00	0.73	0.00	0.02	0.15	0.01	0.08	0.01
class 3	0.01		0.94		0.05	0.00		
class 4				0.93	0.01	0.02	0.03	0.00
class 5	0.01	0.02	0.01	0.02	0.89	0.02	0.03	
class 6	0.02		0.00	0.01	0.01	0.94	0.02	0.00
class 7	0.00	0.01		0.02	0.02	0.03	0.91	0.00
class 8	0.01			0.00		0.01	0.00	0.98

(b)

	class 1	class 2	class 3	class 4	class 5	class 6	class 7	class 8
class 1	0.98		0.00		0.00	0.02		
class 2	0.00	0.78	0.00	0.01	0.13	0.03	0.05	0.00
class 3	0.03	0.00	0.95		0.01	0.01		
class 4		0.00		0.91	0.01	0.05	0.02	0.01
class 5	0.02	0.02	0.01	0.01	0.87	0.04	0.02	
class 6	0.01			0.00	0.00	0.98	0.01	
class 7	0.00	0.01		0.01	0.01	0.04	0.93	0.00
class 8	0.00			0.00		0.00	0.00	0.99

(c)

	class 1	class 2	class 3	class 4	class 5	class 6	class 7	class 8
class 1	0.99				0.00	0.00		
class 2	0.00	0.84		0.00	0.10	0.01	0.04	
class 3	0.01		0.97		0.01			
class 4				0.96	0.01	0.02	0.01	
class 5	0.01	0.01	0.00	0.01	0.94	0.01	0.02	
class 6	0.02			0.00	0.00	0.97	0.01	
class 7	0.00	0.01		0.01	0.01	0.01	0.96	
class 8								1.00

(d)

Fig. 12. Confusion matrices of classification maps obtained by (a) 1-D CNN, (b) 2-D CNN, (c) 3-D CNN, and (d) 1D–2D CNN. Class 1: Jurassic sandstone and slate; Class 2: Early Paleozoic marble; Class 3: Triassic sandstone and slate; Class 4: Cambrian granitic gneiss; Class 5: Paleozoic biotite quartz schist; Class 6: Quaternary; and Class 7: Himalaya Leucogranite.

information by 3-D and 1D–2D CNNs improves approximately 5%–10% classification accuracy than using either spatial or spectral features. The execution time reflects the complexity of CNN model from another side. It is obvious that 1-D CNN costs less time due to the simplest network structure and least parameters. The 1D–2D CNN has a more complex network architecture, and unexpectedly, it performs more efficiently than 3-D CNN, indicating that 3-D convolution occupies more computational burden and parameters than 2-D convolution. The detailed analysis of the classification performance related to each lithological unit is summarized as follows.

Jurassic sandstone and slate (Class 1): For the Jurassic sandstone and slate with the largest areas, each model predicts high and unified classification accuracy (>96%). The main reason is that these models were completely trained with sufficient samples.

Early Paleozoic marble (Class 2): The Early Paleozoic marble exposes the smallest areas compared with other classes, leading to poor performance for all models. Using only spectral information causes the 1-D CNN to fail to extract more discriminative features, whereas adding spatial information from adjacent pixels significantly enhances the classification accuracy.

Triassic sandstone and slate (Class 3): Triassic sandstone and slate occupy independent parts of the study area. The 1-D CNN correctly identified 66% of this lithological unit, while most of them are misclassified as Jurassic sandstone and slate (Class 1). The indistinguishable spectral curves of these two classes caused by similar mineral compositions are regarded as the main limitations. The extraction of spatial information by 2-D CNN improved the separability between various classes, increasing the classification accuracy by 40%. Furthermore, the integration of spatial–spectral features facilitated a classification accuracy of over 97%, significantly reducing the misclassification of highly similar lithological units. This finding implies that spatial features play essential roles in identifying this type of lithological unit.

Cambrian granitic gneiss (Class 4): The Cambrian granitic gneiss is scattered throughout the core of the study area. It can be observed from the confusion matrix that the spectral information is unhelpful in improving the classification accuracy for this unit. The 2-D CNN wrongly classifies a part of the Cambrian granitic gneiss into the early Paleozoic marble (Class 2); however, the 3-D and 1D–2D CNNs correct this error, suggesting that the recognition of Cambrian granitic gneiss mainly depends on texture features.

Paleozoic biotite quartz schist (Class 5): The Paleozoic biotite quartz schist spreads over the periphery of the study area in strip form. A broad misclassification occurred in the four networks. A reasonable explanation for this lies in the extensive contact with other lithological units, resulting in similar lithologies caused by intense metamorphism. Mixed spectral curves make it more difficult to distinguish between these lithological units. Similar to the Cambrian granitic gneiss (Class 4), spectral information brings no contribution for improving the classification performance. On the contrary, the fusion of spectral features instead reduces classification accuracy, indicating from the side that 3-D CNN is more suitable for those lithological classes with regular shapes. Fortunately, the joint network of 1-D and 2-D CNNs yields a satisfactory overall accuracy higher than 93%, which implies that 1D–2D CNN can distinguish subtle differences in highly similar lithological units through deep-level networks, rather than focusing on simple features, such as spectra or textures.

Quaternary (Class 6): The Quaternary strata are mainly distributed in the center of the study area, which are in contact with all other lithological units. Extracting either spectral or spatial information is meaningless for improving classification accuracy. However, extracting both spatial and spectral information can enhance classification performance to a certain extent. The 1D–2D CNN yields 3-D CNN by wrongly dividing part of the Quaternary into Jurassic sandstone and slate (Class 1). This is more likely due to the correlated spectral curves caused by weathering and erosion.

Himalayan leucogranite (Class 7): The Himalayan leucogranite is considered the primary target for exploring rare metal deposits in this area. Satisfactory identifications were realized by all four networks. However, some leucogranites are still misclassified as other classes. These lithological units belong to silicates that are beyond the diagnosis of GF-5 hyperspectral image. The 2-D CNN achieved 90.79% classification accuracy, and 3-D CNN earned a 2% improvement with the support of spatial–spectral information. The mixed spectral curves and indistinguishable spatial texture caused by similar geochemical and geophysical characteristics of adjacent lithological units are the main reasons for this misclassification. Nevertheless, a classification accuracy of up to 96.4% was sufficient to demonstrate the potential application of 1D–2D CNN in lithological mapping.

Hyperspectral remote sensing image combines high-resolution spectral bands with spatial information that represent the distribution and texture features, enabling to discriminate closely resembling minerals for fine lithological mapping. CNNs can automatically extract deep-level spatial–spectral features through layer-by-layer convolution operations, providing a way for handling high dimensionality and mixed pixels of hyperspectral data. Case studies in this article illustrated that the proposed dual-branch 1D–2D network achieved a marked improvement in both visual and quantitative comparisons. The 1D–2D CNN captures more training data to mine deeper prior knowledge, thereby greatly enhancing the discrimination of spectrally closed minerals. Regarding the misclassification of lithological units, this might be addressed by conducting mineral spectral analysis in fieldwork and laboratories, in conjunction with fusing more multisource data, such as thermal infrared images, geochemical, and geophysical information.

However, although CNNs have evolved extensively in geological applications, there are still several challenges raised. The 1-D CNN fails to extract spatial and texture features, resulting in noisy result when applied to lithological mapping; 2-D CNN considers local spatial information around each pixel; however, 2-D CNN is not suitable for lithological mapping with irregular shapes due to fixed convolution kernels; 3-D CNN enables to learn spectral–spatial features simultaneously, significantly improving the classification accuracy, which comes with computational costs as a result of the increased number of parameters. In addition, multiscale and multilevel feature fusion strategies [42] can be employed to deeply integrate spectral–spatial information to improve classification performance and time–cost, which deserves further study. In a word, this study provides a view of lithological mapping to support mineral exploration using hyperspectral images and deep learning methods. The proposed workflow is worthy of promotion for geological mapping work under harsh environments and is applicable to a variety of lithologies or minerals mapping.

VII. CONCLUSION

Hyperspectral images provide both spectral and spatial information that occupy an important position in remote sensing geology, such as geological mapping and mineral exploration. Recent advances in deep learning techniques have facilitated the application of hyperspectral images from qualitative discrimination to quantitative identification. On this account, this study focuses on the evaluation of CNNs to extract spectral and spatial features from hyperspectral data for lithological mapping and further provides guidance on mineral prospectivity mapping for rare metal deposits. The main conclusions of this study are as follows.

1) Four deep learning networks, 1-D CNN, 2-D CNN, 3-D CNN, and 1D–2D CNN, were compared for lithological mapping based on hyperspectral images. The 1-D and 2-D CNNs consider spectral or spatial information of hyperspectral data, respectively, whereas the 3-D and 1D–2D CNNs can learn spatial and spectral information simultaneously.
2) The four CNNs were illustrated by mapping seven lithological units in the Cuonadong dome using GF-5 hyperspectral images, which have been confirmed to be closely related to rare metal mineralization. Classification maps with a maximum overall accuracy of up to 97.4% were achieved, indicating the potential value of CNN models for lithological mapping.
3) The 1D–2D CNN yields better results in both visual and quantitative comparisons. Satisfactory performance benefits from the joint extraction of spatial and spectral features, providing a realizable and valuable approach for mineral exploration by mapping specific lithological units.

COMPUTER CODE AVAILABILITY

The detailed codes used in this study are provided in GitHub (https://github.com/ziyewangcug/CNNs for lithological mapping.git).

REFERENCES

[1] M. Borengasser, W. S. Hungate, and R. Watkins, *Hyperspectral Remote Sensing: Principles and Applications*. Boca Raton, FL, USA: CRC Press, 2007.

[2] A. Khan, A. D. Vibhute, S. Mali, and C. H. Patil, “A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications,” *Ecol. Inform.*, vol. 69, 2022, Art. no. 101678.

[3] M. Shimoni, R. Haelterman, and C. Perneel, “Hyperspectral imaging for military and security applications: Combining myriad processing and sensing techniques,” *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 101–117, Jun. 2019.

[4] S. Veraverbeke et al., “Hyperspectral remote sensing of fire: State-of-the-art and future perspectives,” *Remote Sens. Environ.*, vol. 216, pp. 105–121, 2018.

[5] F. D. van der Meer et al., “Multi- and hyperspectral geologic remote sensing: A review,” *Int. J. Appl. Earth Observ. Geoinf.*, vol. 14, no. 1, pp. 112–128, 2012.

[6] S. Peighambari and Y. Zhang, “Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: An updated review,” *J. Appl. Remote Sens.*, vol. 15, no. 3, pp. 1–25, 2021.

[7] X. Dong, F. Gan, N. Li, S. Zhang, and T. Li, “Mineral mapping in the Duolong porphyry and epithermal ore district, Tibet, using the Gaofen-5 satellite hyperspectral remote sensing data,” *Ore Geol. Rev.*, vol. 151, 2022, Art. no. 105222.

[8] G. R. Hunt, “Spectroscopic properties of rocks and minerals,” in *Handbook of Physical Properties of Rocks*, vol. 1. Boca Raton, FL, USA: CRC Press, 2017, pp. 295–385.

[9] E. A. Cloutis, “Review article hyperspectral geological remote sensing: Evaluation of analytical techniques,” *Int. J. Remote Sens.*, vol. 17, no. 12, pp. 2215–2242, 1996.

[10] R. N. Clark et al., “USGS digital spectral library splib06a,” US Geol. Surv., Reston, VA, USA, Open File Rep., 2003.

[11] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, “Deep learning for hyperspectral image classification: An overview,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6690–6709, Sep. 2019.

[12] L. He, J. Li, C. Liu, and S. Li, “Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1579–1597, Mar. 2018.

[13] D. Ramakrishnan and R. Bharti, “Hyperspectral remote sensing and geological applications,” *Curr. Sci.*, vol. 108, no. 5, pp. 879–891, 2015.

[14] M. Pal, T. Rasmussen, and A. Porwal, “Optimized lithological mapping from multispectral and hyperspectral remote sensing images using fused multi-classifiers,” *Remote Sens.*, vol. 12, no. 1, 2020, Art. no. 177.

[15] B. Kumar, O. Dikshit, A. Gupta, and M. K. Singh, “Feature extraction for hyperspectral image classification: A review,” *Int. J. Remote Sens.*, vol. 41, no. 16, pp. 6248–6287, 2020.

[16] C. Rodarmel and J. Shan, “Principal component analysis for hyperspectral image classification,” *Surveying Land Inf. Sci.*, vol. 62, no. 2, pp. 115–122, 2002.

[17] J. Wang and C.-I. Chang, “Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1586–1600, Jun. 2006.

[18] N. Xu, Y. Hu, B. Lei, Y. Hong, and F. Dang, “Mineral information extraction for hyperspectral image based on modified spectral feature fitting algorithm,” *Spectrosc. Spectral Anal.*, vol. 31, no. 6, pp. 1639–1643, 2011.

[19] J. R. Harris, D. Rogge, R. Hitchcock, O. Ijewliw, and D. Wright, “Mapping lithology in Canada’s Arctic: Application of hyperspectral data using the minimum noise fraction transformation and matched filtering,” *Can. J. Earth Sci.*, vol. 42, no. 12, pp. 2173–2193, 2005.

[20] D. Lunga, S. Prasad, M. M. Crawford, and O. Ersoy, “Manifold-learning-based feature extraction for classification of hyperspectral data: A review of advances in manifold learning,” *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 55–66, Jan. 2014.

[21] T. V. Bandos, L. Bruzzone, and G. Camps-Valls, “Classification of hyperspectral images with regularized linear discriminant analysis,” *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 862–873, Mar. 2009.

[22] Y. Tan, L. Lu, L. Bruzzone, R. Guan, Z. Chang, and C. Yang, “Hyperspectral band selection for lithologic discrimination and geological mapping,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 471–486, 2020.

[23] Y. Dong, Y. Jin, and S. Cheng, “Clustered multiple manifold metric learning for hyperspectral image dimensionality reduction and classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5516813.

[24] Z. Ren, L. Sun, and Q. Zhai, “Improved k-means and spectral matching for hyperspectral mineral mapping,” *Int. J. Appl. Earth Observ. Geoinf.*, vol. 91, 2020, Art. no. 102154.

[25] X. Zhang and P. Li, “Lithological mapping from hyperspectral data by improved use of spectral angle mapper,” *Int. J. Appl. Earth Observ. Geoinf.*, vol. 31, pp. 95–109, 2014.

[26] F. Melgani and L. Bruzzone, “Classification of hyperspectral remote sensing images with support vector machines,” *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[27] J. Ham, Y. Chen, M. M. Crawford, and J. Ghosh, “Investigation of the random forest framework for classification of hyperspectral data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 492–501, Mar. 2005.

[28] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, “CoSpace: Common subspace learning from hyperspectral-multispectral correspondences,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4349–4359, Jul. 2019.

[29] R. Guan, M. Wang, L. Bruzzone, H. Zhao, and C. Yang, “Lightweight attention network for very high resolution image semantic segmentation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 61, May 2023.

[30] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, “Graph convolutional networks for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021.

[31] L. Mou, P. Ghamisi, and X. X. Zhu, “Deep recurrent neural networks for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639–3655, Jul. 2017.

[32] P. Zhong, Z. Gong, S. Li, and C.-B. Schönlieb, “Learning to diversify deep belief networks for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3516–3530, Jun. 2017.

[33] P. Zhou, J. Han, G. Cheng, and B. Zhang, “Learning compact and discriminative stacked autoencoder for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4823–4833, Jul. 2019.

[34] L. Zhu, Y. Chen, P. Ghamisi, and J. A. Benediktsson, “Generative adversarial networks for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5046–5063, Sep. 2018.

[35] D. Datta, P. K. Mallick, A. K. Bhoi, M. F. Ijaz, J. Shafi, and J. Choi, “Hyperspectral image classification: Potentials, challenges, and future directions,” *Comput. Intell. Neurosci.*, vol. 2022, 2022, Art. no. 3854635.

[36] B. Zhang, S. Li, X. Jia, L. Gao, and M. Peng, “Adaptive Markov random field approach for classification of hyperspectral imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 973–977, Sep. 2011.

[37] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, “Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles,” *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.

[38] X. Huang, X. Liu, and L. Zhang, “A multichannel gray level co-occurrence matrix for multi/hyperspectral image texture representation,” *Remote Sens.*, vol. 6, no. 9, pp. 8424–8445, 2014.

[39] S. Jia, K. Wu, J. Zhu, and X. Jia, “Spectral–spatial Gabor surface feature fusion approach for hyperspectral imagery classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1142–1154, Feb. 2019.

[40] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, “Advances in spectral-spatial classification of hyperspectral images,” *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[41] M. Ahmad et al., “Hyperspectral image classification—Traditional to deep models: A survey for future prospects,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 968–999, Dec. 2021.

[42] M. Imani and H. Ghassemian, “An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges,” *Inf. Fusion*, vol. 59, pp. 59–83, 2020.

[43] R. Grewal, S. Singh Kasana, and G. Kasana, “Machine learning and deep learning techniques for spectral spatial classification of hyperspectral images: A comprehensive survey,” *Electronics*, vol. 12, no. 3, 2023, Art. no. 488.

[44] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, “HybridSN: Exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification,” *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 277–281, Feb. 2020.

[45] F. Viel, R. C. Maciel, L. O. Seman, C. A. Zeferino, E. A. Bezerra, and V. R. Q. Leithardt, “Hyperspectral image classification: An analysis employing CNN, LSTM, transformer, and attention mechanism,” *IEEE Access*, vol. 11, pp. 24835–24850, 2023.

[46] Y. Li, H. Zhang, and Q. Shen, “Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network,” *Remote Sens.*, vol. 9, no. 1, 2017, Art. no. 67.

[47] Y. Xu, B. Du, and L. Zhang, “Beyond the patchwise classification: Spectral-spatial fully convolutional networks for hyperspectral image classification,” *IEEE Trans. Big Data*, vol. 6, no. 3, pp. 492–506, Sep. 2020.

[48] K. Zhu, Y. Chen, P. Ghamisi, X. Jia, and J. A. Benediktsson, “Deep convolutional capsule network for hyperspectral image spectral and spectral-spatial classification,” *Remote Sens.*, vol. 11, no. 3, 2019, Art. no. 223.

[49] Y. Xu, L. Zhang, B. Du, and F. Zhang, “Spectral–spatial unified networks for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5893–5909, Oct. 2018.

[50] W.-S. Hu, H.-C. Li, L. Pan, W. Li, R. Tao, and Q. Du, “Spatial–spectral feature extraction via deep ConvLSTM neural networks for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 4237–4250, Jun. 2020.

[51] H. Zhang, Y. Li, Y. Zhang, and Q. Shen, “Spectral-spatial classification of hyperspectral imagery using a dual-channel convolutional neural network,” *Remote Sens. Lett.*, vol. 8, no. 5, pp. 438–447, 2017.

[52] M. R. Haque, S. Z. Mishu, P. Uddin, and M. Al Mamun, “A lightweight 3D-2D convolutional neural network for spectral-spatial classification of hyperspectral images,” *J. Intell. Fuzzy Syst.*, vol. 43, no. 1, pp. 1241–1258, 2022.

[53] S. Ghaderizadeh, D. Abbasi-Moghadam, A. Sharifi, N. Zhao, and A. Tariq, “Hyperspectral image classification using a hybrid 3D-2D convolutional neural networks,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 7570–7588, Jul. 2021.

[54] S. Mei, J. Ji, Y. Geng, Z. Zhang, X. Li, and Q. Du, “Unsupervised spatial–spectral feature learning by 3D convolutional autoencoder for hyperspectral classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6808–6820, Sep. 2019.

[55] A. Qin, Z. Shang, J. Tian, Y. Wang, T. Zhang, and Y. Y. Tang, “Spectral–spatial graph convolutional networks for semisupervised hyperspectral image classification,” *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 241–245, Feb. 2019.

[56] W. Zhao and S. Du, “Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544–4554, Aug. 2016.

[57] S. Li, X. Zhu, Y. Liu, and J. Bao, “Adaptive spatial-spectral feature learning for hyperspectral image classification,” *IEEE Access*, vol. 7, pp. 61534–61547, 2019.

[58] J. Yang, Y.-Q. Zhao, and J. C.-W. Chan, “Learning and transferring deep joint spectral–spatial features for hyperspectral classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4729–4742, Aug. 2017.

[59] H. Sun, X. Zheng, X. Lu, and S. Wu, “Spectral–spatial attention network for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3232–3245, May 2020.

[60] B. Ye, S. Tian, Q. Cheng, and Y. Ge, “Application of lithological mapping based on advanced hyperspectral imager (AHSI) imagery onboard Gaofen-5 (GF-5) satellite,” *Remote Sens.*, vol. 12, no. 23, 2020, Art. no. 3990.

[61] H. Shirmard et al., “A comparative study of convolutional neural networks and conventional machine learning models for lithological mapping using remote sensing data,” *Remote Sens.*, vol. 14, no. 4, 2022, Art. no. 819.

[62] H. Shirmard, E. Farahbakhsh, R. D. Müller, and R. Chandra, “A review of machine learning in processing remote sensing data for mineral exploration,” *Remote Sens. Environ.*, vol. 268, 2022, Art. no. 112750.

[63] C. Zhang, M. Yi, F. Ye, Q. Xu, X. Li, and Q. Gan, “Application and evaluation of deep neural networks for airborne hyperspectral remote sensing mineral mapping: A case study of the Baiyanghe uranium deposit in Northwestern Xinjiang, China,” *Remote Sens.*, vol. 14, no. 20, 2022, Art. no. 5122.

[64] G. M. Li et al., “First discovery and implications of Cuonadong superlarge Be-W-Sn polymetallic deposit in Himalayan metallogenic belt, southern Tibet,” *Mineral Deposits*, vol. 36, no. 4, pp. 1003–1008, 2017.

[65] F. Wu et al., “Highly fractionated Himalayan leucogranites and associated rare-metal mineralization,” *Lithos*, vol. 352, 2020, Art. no. 105319.

[66] C.-T. He et al., “Multiple skarn generations related to composite leucogranites in the Cuonadong Sn-W-Be deposit, Himalaya,” *Ore Geol. Rev.*, vol. 150, 2022, Art. no. 105161.

[67] H.-W. Cao et al., “Himalayan leucogranites: A review of geochemical and isotopic characteristics, timing of formation, genesis, and rare metal mineralization,” *Earth-Sci. Rev.*, vol. 234, 2022, Art. no. 104229.

[68] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[69] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[70] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, vol. 25.

[71] J. Gu et al., “Recent advances in convolutional neural networks,” *Pattern Recognit.*, vol. 77, pp. 354–377, 2018.

[72] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 3, no. 12, pp. 6999–7019, Dec. 2022.

[73] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mech. Syst. Signal Process.*, vol. 151, 2021, Art. no. 107398.

[74] Z. Ge, G. Cao, X. Li, and P. Fu, "Hyperspectral image classification method based on 2D–3D CNN and multibranch feature fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5776–5788, 2020.

[75] M. Ahmad, A. M. Khan, M. Mazzara, S. Distefano, M. Ali, and M. S. Sarfraz, "A fast and compact 3-D CNN for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, Dec. 2020.

[76] J. Yang, Y. Zhao, J. C. Chan, and C. Yi, "Hyperspectral image classification using two-channel deep convolutional neural network," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2016, pp. 5079–5082.

[77] H. C. Kraemer, "Kappa coefficient," in *Wiley StatsRef: Statistics Reference Online*. Hoboken, NJ, USA: Wiley, 2014, pp. 1–4.

[78] D. Wang et al., "Multiple mineralization events in the Zhaxikang Sb–Pb–Zn–Ag deposit and their relationship with the geodynamic evolution in the North Himalayan Metallogenic Belt, South Tibet," *Ore Geol. Rev.*, vol. 105, pp. 201–215, 2019.

[79] J. Fu et al., "First field identification of the Cuonadong dome in southern Tibet: Implications for EW extension of the North Himalayan gneiss dome," *Int. J. Earth Sci.*, vol. 106, pp. 1581–1596, 2017.

[80] H. Cao et al., "Miocene Sn polymetallic mineralization in the Tethyan Himalaya, southeastern Tibet: A case study of the Cuonadong deposit," *Ore Geol. Rev.*, vol. 119, 2020, Art. no. 103403.

[81] H. Cao et al., "Genesis of the Cuonadong tin polymetallic deposit in the Tethyan Himalaya: Evidence from geology, geochronology, fluid inclusions and multiple isotopes," *Gondwana Res.*, vol. 92, pp. 72–101, 2021.

[82] X. B. Xia, G. M. Li, H. W. Cao, W. Liang, and J. G. Fu, "Petrogenic age and geochemical characteristics of the mother rock of skarn type ore body in the Cuonadong Be-W-Sn polymetallic deposit, Southern Tibet," *Earth Sci.*, vol. 44, no. 7, pp. 2207–2223, 2019.

[83] L. Gao, L. Zeng, and P. D. Asimow, "Contrasting geochemical signatures of fluid-absent versus fluid-fluxed melting of muscovite in metasedimentary sources: The Himalayan leucogranites," *Geology*, vol. 45, no. 1, pp. 39–42, 2017.

[84] R. R. Girija and S. Mayappan, "Mapping of mineral resources and lithological units: A review of remote sensing techniques," *Int. J. Image Data Fusion*, vol. 10, no. 2, pp. 79–106, 2019, doi: 10.1080/19479832.2019.1589585.

[85] Y.-N. Liu et al., "The advanced hyperspectral imager: Aboard China's gaoFen-5 satellite," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 4, pp. 23–32, Dec. 2019.

[86] T. Cooley et al., "FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2002, vol. 3, pp. 1414–1418.

[87] C. P. Dave, R. Joshi, and S. S. Srivastava, "A survey on geometric correction of satellite imagery," *Int. J. Comput. Appl.*, vol. 116, no. 12, pp. 24–27, 2015.

[88] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.

[89] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, 2020.

[90] M. Feurer and F. Hutter, "Hyperparameter optimization," in *Automated Machine Learning: Methods, Systems, Challenges*. Cham, Switzerland: Springer, 2019, pp. 3–33.

[91] R. Meyes, M. Lu, C. W. de Puiseau, and T. Meisen, "Ablation studies in artificial neural networks," 2019, *arXiv:1901.086A44*.

Ziye Wang received the B.S. degree in mining engineering from the Henan University of Technology, Jiaozuo, China, in 2014, and the M.S. and Ph.D. degrees in geo-detection and information technology from the China University of Geosciences (CUG), Wuhan, China, in 2017 and 2020, respectively.

He has authored or coauthored more than 20 papers. He is currently an Associate Professor with the School of Earth Resources, CUG. His research interests include remote sensing geology and mathematical geosciences.

Renguang Zuo received the B.S. in resource exploration engineering and the Ph.D. in mineral prospecting and exploration degree from the China University of Geosciences (CUG), Wuhan, China, in 2004 and 2009, respectively.

He is currently a Full Professor with the State Key Laboratory of Geological Processes and Mineral Resources, CUG. His current research interests include mathematical geosciences and mineral exploration. He has authored or coauthored more than 120 international peer-reviewed papers in various journals. He is an Associate Editor for *Computers and Geosciences*, *Journal of Geochemical Exploration*, *Natural Resources Research*, and *Ore Geology Reviews*.